December 2002

C-ASSURE-TAGUCHI FRAMEWORK FOR COST-EFFECTIVE HOLISTIC HEURISTIC IS EVALUATIONS

Atta Badii
University College Northampton

Andy Smith
University of Luton

C-ASSURE-TAGUCHI FRAMEWORK FOR COST-EFFECTIVE HOLISTIC HEURISTIC IS EVALUATIONS

Atta Badii
P3ie-EnCKompass Research Group
Information Systems Department
University College Northampton
atta.badii@northampton.ac.uk

Andy Smith
Department of Computing
University of Luton
Andy.smith@luton.ac.uk

Abstract

Aspiring to achieve holistic IS evaluations is fraught with data acquisition and interpretation problems arising from the multi-faceted nature of empirical data on all aspects of usability, actability and user preferences. The relevant data consists of both quantitative and qualitative elements arising from both the requirement specification and design parameter optimisation phase as well as the IS deployment impact elicitation and analysis phase. Some of this data may be capable of interpretation or processing using parametric statistics or dynamic optimisation techniques whilst other data elements, for example organisational impact data, will only lend themselves to the narrative of situated interpretation and local sense making. C-Assure has been previously applied to the elicitation, evaluation and thus optimisation of direct and indirect socio-technical consequences of IS deployment in a distributed actor environment whilst the Taguchi method had been powerfully applied to support the evaluation and thus optimisation of alternative design specifications of advanced information systems, for example Web Portals.

In this paper the authors examine the integration of the above two complementary and cost-effective approaches to information systems (IS) evaluation analysis namely the C-Assure and Taguchi methods. To illustrate the efficacy of the resulting C-Assure-Taguchi platform we first examine the potential for additional usability optimisation, particularly dynamic usability theory and man-machine relationship understanding using C-Assure (Badii 2000a,b). This is applied to the cost-effective but holistic heuristic usability and impact analysis of an enterprise information portal for the insurance industry. Additionally the Taguchi method as applicable to IS design optimality analysis (Smith and Dunkley 1998) is deployed to conclude the ranking of some of the key factors for optimisation of the lower level design parameters of the above portal. In this study these factors were found to be usability, branding and speed in ranked order of their influences re the overall impact of the website on its typical target users. In future studies it is planned to deconstruct design usability further into sets of drivers that are as linearly independent as possible thus assuring optimisation over the key orthogonal variable space i.e. best global trade-off of design parameters to support user satisfaction.

Further we conclude that the above approaches can contribute much to the effective and efficient usability evaluation and optimisation of information systems and in particular web-mediated systems and may be used separately or together as part of a Web Systems and Services design toolkit of methods and guidelines for Holistic Heuristic (H^2) Usability-Actability Evaluation and impact analysis.

Introduction

Systems Success or Failure?

Figure 1. Introducing IS into the Organisation

Several researchers (e.g. Rowland 2000, Seminario 1998) have pointed out the crucial impact that web usability can have on competitive advantage particularly in dominantly e-channel oriented value-chains. In recent years many researchers have analysed the phenomenology associated with the failures in deployment of information systems into organisations; (e.g. Hirschheim & Smithson 1999) particularly in the government sector. Nearly a decade ago when discussing the 'chronic' crisis of software usability, Gibbs (1994) stated that three quarters of all large systems were operating failures that either did not function as intended, or were not used at all. Even earlier research (e.g. Butler Cox Foundation 1986, Lyytinen 1988) suggests that up to half of information systems projects when introduced into the social system of the organisation (as depicted in Figure 1) fail to become successful or yield the expected return on investments.

Concurrently a number of researchers (e.g. Eason 1988) have emphasised the importance of the *socio-technical* nature of information systems, and the need to provide not just a technical systems product (hardware and software), but also a social systems product (organisation and job design), within a user-centred approach to systems development (Smith 1997). Furthermore it is now generally understood that the impact of the information system on its end users is critical in determining success and, particularly, in promoting a healthy ecology of patterns-of-power-and-relating within the actors' environment (Badii 2000a,b). The latter, we maintain, crucially determines the prospect for organisational resilience and enduring success. However in the context of web-mediated man-machine relationship building, new theoretical frontiers have been opened by dynamic usability theory advances as incorporated in C-Assure (Badii 2000). Essentially C-Assure aids modelling and reasoning about the likely cultural match and usability impact resulting from IS deployment over the whole lifecycle. This extends from inception to development into the environment through to ongoing post deployment usability relationships situation assessment. This paper reviews these new approaches to the user centric evaluation of information systems and in particular proposes the incorporation of such methods as important elements of an IS usability evaluation toolkit whereby the crucial drivers of usability and technology acceptance can be identified and measured.

IS Impact Evaluation

It may be useful to start this analysis by looking at what we mean by an information system. Gupta (1996) states that an information system 'is a system that creates, processes, stores and retrieves information'. Whilst it would be difficult to disagree with this definition we need to explore the full entailments of an information system. We are firmly in agreement with Checkland (1991) and others that the thinking of an IS as starting from a means (a computer) is flawed and that one that addresses its ends (an organisation's conceptualisation of its world) is more appropriate. In essence we need to treat IS more as a cultural rather than a technical phenomenon. IS culture can itself exist at different levels from the organisational level to the individual level. At the individual level the success of IS is often referred to in terms of user acceptance or rejection. Herschheim and Newman (1988) provide an extensive review of the theory and practice of user resistance to the introduction of computer systems. Eason (1998) describes a method by which the impact of a proposed technical system on a target organisation can be judged.

The process, as depicted in a slightly modified form in Figure 2, identified four stages:

- User group identification, through which the individual user groups that could possibly be affected by the proposed technical system are identified
- User group cost benefit analysis, in which the potential effects of the proposal on each user group can be identified
- User reaction assessment, through which the probable responses of both winners and losers in the organisation are identified
- Organisational cost benefit analysis, in which the global effect on the organisation and user groups is summarised

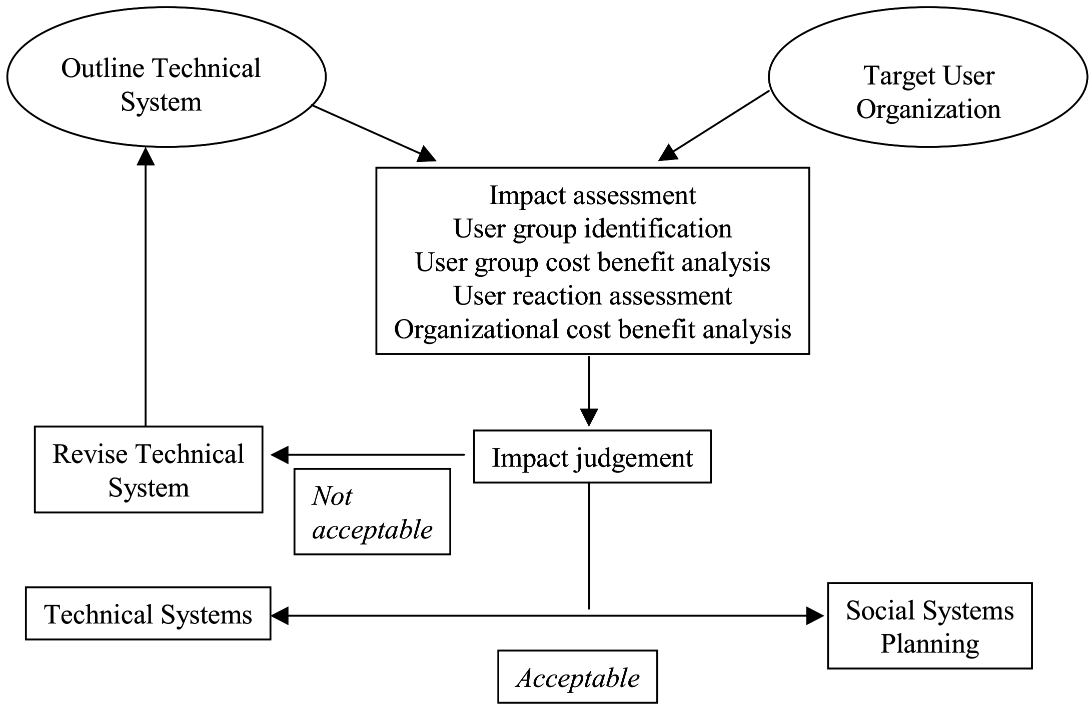

Figure 2. IS Impact Evaluation

As a result of impact assessment an overall judgement can be made as to whether a proposed Information System is acceptable (will be successful) or not. Well-founded criteria are essential to the accurate assessment of the impact of systems on organisations and user groups. The criteria should be based upon the specific ways in which the IS may affect individuals in each user group.

User Impact Evaluation

This part of the paper focuses on the third level of Eason's method for impact assessment: user reaction assessment, and in particular investigates two approaches whereby the characteristics inherent within Information Systems that affect usability and impact user relationship building can be both measured and rank ordered. What we are trying to do here is to find out, through user impact assessment studies, not just whether the IS will be acceptable / successful, but why this is so. Such reflectivity applied to the design of the IS-mediated ecology of actors' environment seeks to elicit, and respond to, the issues that are the key drivers of user (dis) satisfaction? i.e. what is the contribution of each element of the organisational actability ecology as mediated by IS?

The C-Assure Methodology, Dynamic Usability Theory and Actability

Here we briefly outline the principles of Dynamic Usability Theory and Human Judgement and Decision Making (J/DM) theory as well as theories of Pleasure and Pain Recall (PPR) integrated into those components of C-Assure that relate to the social psychology of computing (Badii 2000a,b). Holistic Evaluation and its C-Assure-mediated enhancement are then described followed by a discussion of results and future work. An important element of this framework is the development of the multi-dimensional Effects-Affects Usability Matrices. These link C-Assure *effects*, *side-effects* and *affect*s on people, processes, partners, places and portals with C-Assure-theoretic models of perceived comfort, pleasure, pain and preferences (Badii 2000a, b). The paper concludes by giving an outline of results of the C-Assure-mediated heuristic evaluation of the Insurance Claim Settlement Demonstrator and an assessment of the improvements achieved over the standard heuristic evaluation by integrating the C-Assure framework with it. This enhancement is recommended particularly to designers of Web-Mediated Systems (WMS), content architects and persona technologists as an alternative means of conducting a holistic but cost-effective heuristic evaluation of WMSs. Thus we ensure savings in time and cost as well as the added advantages of a richer and more socio-technically-aware methodology by which to assess the efficacy of tactical and strategic IS developments and particularly WMSs.

The primary objective of an interactive system is to allow the user to achieve particular goals in an application domain i.e. the interactive system must be usable (Dix et al 1993). Usability is defined as:

A) Classic Usability: A system is usable if it incorporates learnability, flexibility and robustness (Dix 1993c).
B) Dynamic Usability: Badii (1999-2001) added three new criteria to the classic definition of usability:
- **i)** Mutual Man-Machine Intelligibility and Scanability (M^*IS).
- **ii)** Mutually Productive balance of Power, Privacy and Patterns-of-Relating (MP^*oR) between the man and machine *personas*, as transactors (Badii 1996--2001)
- **iii)** Sustainability of both M^*IS and MP^*oR i.e. SMP^*.

This dynamic usability expands on the classic definition to provide a view of instantaneous usability, which is necessary in dynamic process-centred analysis. It aids formulation and testing of spatio-temporal-causal models of usability by being more expressive of the complex socio-technical dynamics of situated usability. Further the SMP* criterion allows it to explain the effects of such J/DM-PPR theoretic saliency-recency on instantaneous and steady state usability (Badii 2000a,b). Thus this Dynamic Usability is sufficiently richly defined to cope with the increasingly volatile *click-happy user* environment that pertains to Web-Mediated Systems; for example in on-line shopping (Badii 2000a, b). It is important for manufacturers, users, organisations and their change managers to have a facility for developing a high *Perceived usability* or Comfort Factor (PCF) for interactive systems. Several types of, and approaches to, evaluation exist as addressed by many researchers (e.g. Draper & Brown 1996). Our primary focus in this paper is on expert heuristic evaluations that seek to capture the fullest possible holistic picture of usability of a system most cost-effectively.

The important research question for HCI and usability as targeted by previous research (Badii 2000a,b) was: can J/DM-PPR effects be successfully exploited in various cultural contexts in organisations as well as in the attendant change management lifecycle? To investigate the above hypotheses, a methodology of simulation and testing was conducted (Badii 2000a,b) through an empirically grounded study involving multi-modal and multimedia data capture and triangulation of results. This used the bias-minimised on-line usability data capture systems referred to as the *PopEval Family* (PopEval_MB, PopEval_AB) in the context of both Web and non-Web application domains in parallel usability experiments to test and confirm the most important dynamic usability relations as follows:

R_0: Stated that J/DM-PPR bias effects as proven in other domains are also applicable to the usability evaluations of interactive software and in particular to WMSs usability.

R_1: Our Dynamic Usability lends itself to a more readily computable analysis of transitional usability, which can exploit the J/DM-PPR saliency-recency type effects. Thus our previous work confirmed a dynamic usability process model, which can be formalised as:

$$U\mathbf{e}_{Tn} \propto \mathbf{(MP^*oR)}_{T(n-1)} \propto \mathscr{F}\,\mathbf{[(PCF)}_{T_0},\ldots,\mathbf{(PCF)}_{T_{(n-1)}}\mathbf{]}$$

$$\mathscr{F}\,\mathbf{[PCF)}_{T_0},\ldots,\mathbf{(PCF)}_{T_{(n-1)}}\mathbf{]} \propto \mathscr{F}\,\mathbf{[(SI)}_{T_0},\ldots,\mathbf{(SI)}_{T_{(n-1)}}\mathbf{]}$$

This states that usability as evaluated at any time T_n i.e. Ue_{Tn} is a function $\mathscr{F}$ of the mutual man-machine patterns-of-relating i.e. (MP^*oR) as established at the previous instant $T_{(n-1)}$; this in turn depends on the value of the user's Perceived Comfort Factor (PCF) as a global evaluation over the interval $[T_0, T_{(n-1)}]$ and that such PCF is itself a function of the Salient Impression (SI) as formed, *and remembered*, by the user about the system over the same period (i.e. from the beginning to the previous instant). The non-monotonic function $\mathscr{F}$ as applied here to both the SI and PCF is a non-linear function that maps the J/DM-PPR theoretic global evaluations of a user's affective experiences over the transaction interval (i.e. inclusive of natural biases). This approach to usability evaluation provides a finer-grain usability evaluation process data that is remedially prescriptive, deploying insights from cognitive science about the nature of human memory, Judgement and Decision Making, including non-linear hierarchical planning and choice making by actors, and thus involving both the personal and the social psychology of usability.

According to Badii 2000a,b and consistent with this Dynamic Usability Theory, from the standpoint of any C-Assure audit, the usability of a new Information System (IS) depends critically on two properties:

i) the level of sustainable *accommodation* of the users' important and deeply-valued needs afforded by the new IS through its design features,

ii)_ the expected level of adaptation potential of a user group -reachable without them suffering intolerable cultural stress/distress (reachabilities).

The ideal new IS would aim to reduce the level of man-machine P^*oR imbalance or cultural mis-match in attempting to strike realisable and affordable design trade-offs for the whole situated system. Thus C-Assure should allow access to meta-models, tools, repositories, ontologies, and, local-global practice logics to support mass re-negotiation of users' needs, not confining user involvement to "focus groups" only. This will significantly aid the attainment of a higher Perceived Comfort Factor (PCF) as well as cross-cultural interoperability of the evolving Information Systems.

Heuristic Evaluation

Heuristic evaluation is amongst the easiest methods to learn and results in problem reports that are relatively robust predictors of the difficulties that end-users are likely to experience in using the system. A heuristic is a general guide for an activity; what might be described as a "rule of thumb". For example, the heuristics compiled by Nielsen (1994) included such widely accepted principles of user interface design as "supports recognition rather than recall" and "prevents errors". The method uses multiple evaluators who conduct independent inspections in which they compare interface design features with a list of recognised usability principles or heuristics and evaluate the degree to which a desirable interaction facility fulfils performance expectations or fails them. Thus severity ratings are assigned to failures to deliver on certain usability criteria (heuristics) as expected, or, unexpected irritating side effects that can cause aversive experiences amongst end-users and can lead to a generalised user dissatisfaction and the rejection of the system. In heuristic evaluation, a scoring method is usually applied to reflect the level of performance on each deliverable, the severity ratings and the final scores. In such a heuristic evaluation exercise the inspection reports from multiple independent evaluators are considered together in order to maximise the chances of properly identifying the usability problems that might arise. Studies have found that the use of 3 to 5 evaluators is the minimum that will ensure identification of about 75% of usability problems in a project. The use of many more evaluators will result in only marginal improvements in the rate of detection. Heuristic evaluation has adopted a de-facto standard framework of commonly well-established Usability Evaluation Heuristics or design guidelines that can be interpreted to develop special heuristics for individual evaluations given their particular situated context (Nielsen 1994) as follows:

a) Visibility of system status - The system should always keep users informed about what is going on through appropriate feedback within reasonable time.
b) Match between the system and the real world - The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.
c) User control and freedom - Users often choose system functions by mistake and need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undoes and redo.
d) Consistency and standards - Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.
e) Error prevention - Even better than good error messages, which prevents a problem from occurring in the first place.
f) Recognition rather than recall - Make objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another.
g) Instructions for use of the system should be visible or easily retrievable whenever appropriate.
h) Flexibility and efficiency of use – Accelerators, unseen by the novice user, may speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. i.e. allow personalisation.
k) Aesthetic and minimalist design - Dialogues should not contain information which is irrelevant. Every extra unit of information in a dialogue competes with other relevant units of information and diminishes their relative visibility.
l) Help users recognise, diagnose, and recover from errors - Error messages should be in plain language (no codes), precisely indicate the problem, and suggest the solution.
m) Help and documentation - Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and be concise.

C-Assure-Mediated H^2 Evaluation of Enterprise Information Portals

Two teams of expert evaluators well briefed as to the rank ordering of the user needs were deployed; one team was to carry out a standard heuristic evaluation and the other to perform a Holistic Heuristic (H^2) evaluation. The first team of valuators had to consider the standard key WMS properties including for example:

i) Most critical functionalities serving the website purpose?
ii) Areas of the site triggering the heaviest database interaction?
iii) Most complex aspects of the site execution e.g. CGI, applets, ActiveX components etc.
iv) Types of problems causing the most user dis-satisfaction
v) Most frequently visited and most popular parts/paths through the website
vi) Aspects of the site causing the highest security risks

The second team additionally considered the C-Assure-mediated factors, particularly the dynamic usability effects that call for interpretation of some of the standard usability heuristics and thus lead to additional recommended effects matrices to help the evaluation and enhancement of such Websites. These included the specification of the significance ranking of the relevant C-Assure Effects-Affects Matrices and the contingent severity ratings attributable to them within the heuristic evaluation process. A comparison was then made between the two sets of evaluations to assess the relative performance of the two sets of experts in terms of the number and significance of the faults identified and the specificity of their recommendations within broadly the same time and resource constraints as had been available for the two sets of heuristic evaluations. It was found that the C-Assure-mediated H^2 framework did help point out the cooperation-relationship-centred problems *earlier* in the analysis than did the standard heuristic method; and further, it did so *more specifically* in terms of both detection and recommendations for cost-effective remediation and recovery.

Taguchi Methods and IS Evaluation

In essence what we attempt to do here is to describe how Taguchi methods can be implemented in order to relate *the ways in which information systems differ*, to *the factors which determine user impact evaluation*. This section will review of Taguchi methods and how they have been applied within the IT arena to-date. It will then describe a study of their application to IS, particularly Website evaluation, after which some conclusions and areas for further study will be established

An Outline of Taguchi Methods

Genichi Taguchi in Japan developed Taguchi methods for Total Quality Management and their use has spread to both the USA and Europe. The majority of the applications of Taguchi methods are within production control. Underlying the Taguchi method is the concept that quality is affected by two types of factor: internal or *control* factors (such as materials) which can be controlled easily, and external or environmental *noise* factors (such as maintenance of equipment) which cannot be controlled easily. Within the Taguchi approach the adoption of an experimental design strategy enables aspects of quality management to be pushed back from inspection to design. Thus the Taguchi method provides a framework for design, and, a simplified analysis of results, which makes the factorial design accessible to non-statisticians.

Taguchi techniques have not previously been applied to the evaluation of information systems, although they have been applied to the solution of specific software and hardware issues. Smith and Dunkley (1996) and Smith (1997 have integrated Taguchi methods within the Logical User Centred Interface Design (LUCID) method which aims to optimise the usability of the human-computer interface. Indeed the method has been developed to address both internal *interface design* factors and external *user diversity* factors (Smith and Dunckley 1998, Dunckley, Smith and Howard 1999). The basic concept is for the design team to agree on the quality characteristics for measuring the performance of the product. The team then identifies the input factors in the development of the product, which are considered to influence the quality characteristics of the output. The design of the experiments will be based on the selection of an appropriate orthogonal array. Experiments are then conducted to determine values of the quality characteristic associated with the factor levels determined from the orthogonal array. ANOVA is performed to analyse the results and identify the optimum conditions. In fact it is the ability of this approach to address both control and "noise" factors that enables us to apply the Taguchi methods to information systems evaluation. The internal factors can be considered to relate to characteristics of an information systems themselves, whilst the external factors relate to characteristics of the organisation in which such an IS is to be embedded.

The Basic Taguchi Process

The standard Taguchi procedure (Taguchi 1986) for the analysis of internal (control) factors may be summarised as follows:

Factors and Levels

Firstly the design team has to identify the number of key input factors, and the settings or levels of these factors, which they want to test. For example, it could be decided to investigate an input factor at high, medium and low settings, representing three levels. Frequently just two levels, high and low would be used. Factors are included in the study for the purpose of determining their influence and control upon the most desirable performance. Some factors in the design may influence each other and may not be independent. Temperature and humidity, for example, interact in terms of human comfort level. The Taguchi method enables the study of both input factors, and the suspected *interactions* but this falls outside our remit.

Quality Characteristic

The quality characteristic can be a single criterion such as pressure, temperature, efficiency, hardness, or a combination of several criteria considered integratively as a single index. It is necessary to consider the nature of the performance objectives and whether they conform to a *bigger is better*, *smaller is better* or *nominal is the best* type.

Orthogonal Arrays

Orthogonal arrays are a set of tables devised by Taguchi and are used to determine the minimum number of experiments and their input conditions. Table 1 illustrates the first two orthogonal arrays, the L_4 array which would deal with up to three factors and the L_8 array which can cope with up to 7 factors. The orthogonal arrays are systematically named as $L_A(B^C)$ where A is the maximum number of experiments (rows in the orthogonal array), B is the number of levels and C is the number of design factors (columns in the orthogonal array). When the factors and levels are agreed the orthogonal array can be selected, assigning the factors to columns and determining the conditions for individual experiments.

Table 1. The First Two Orthogonal Arrays

$L_4(2^3)$	Column				$L_8(2^7)$	Column						
Condition	**1**	**2**	**3**		**Condition**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1	1	1	1		1	1	1	1	1	1	1	1
2	2	2	2		2	1	1	1	2	2	2	2
3	2	1	2		3	1	2	2	1	1	2	2
4	2	2	1		4	1	2	2	2	2	1	1
					5	2	1	2	1	2	1	2
					6	2	1	2	2	1	2	1
					7	2	2	1	1	2	2	1
					8	2	2	1	2	1	1	2

Running Experiments

There are two common ways of running experiments. If an experiment uses an L_8 array and each trial is repeated 3 times, there will be 24 experiments. In *replication* mode, which is the desired option, all the 24 experiments should be run in random order. In *repetition* mode, which may be the most practical, the trial condition is selected in random order then all repetitions in that trial are run.

Analysis

The main effects are evaluated, and, in the H^2 mode, their influences are determined both qualitatively and quantitatively as appropriate to give the optimum condition. ANOVA is performed on the relevant sets to identify the relative strengths of the factors. Multiple runs can be carried out, the results analysed, and the 'signal-to-noise ratios' (S/N) calculated. The S/N ratio represents a concept developed by Taguchi to be an estimate of the relative strengths of the system factors compared to the 'noise' factors in the environment. A further test of the optimum conditions must be made to confirm the performance.

Taguchi Website Evaluation

Experimental Design

The study described here was undertaken to ascertain the importance of a range of Website design features to users in their assessment of such systems. In the new media arena it has long been believed that download **speed** has been a critical determinant of Internet trading success. More generally however, **usability** has been seen to be an increasingly important factor. Also in relation to user assessment of websites the degree to which the site appears to be owned by a well-established company ('clicks and mortar') as opposed to an Internet-only operation ('dotcom') can be important. This latter issue can be referred to as **branding.** This essentially includes the effect of mixed web and/or non-web presence and branding.

Of course usability, speed and branding are only three amongst many relevant factors but these were chosen to form the basis of this initial study because they represented the most salient, albeit compound, dimensions as a first proving ground.

In the study a well-known car supermarket site was selected. Usability specialists were asked to analyse the site and specify a number of design changes that would improve the usability of the site. It was therefore possible to build versions of the site with **current** or **improved usability**. In order to investigate the effects of download times it was decided to use optimisation software so that the site could be run at both **slow** and **fast speeds**. Users were recruited from the car supermarkets client list so that they would all be aware of the sites brand. However in order to mimic an unknown brand, versions of the site were built with all references to the real life car supermarket brand changed to a fictitious one.In relation to Taguchi methods the study adopted an L4 design with three factors each operating at two levels as shown in Table 2. In total four versions of the car supermarket website were produced, with Site 1 (the first row in Table 2) being the currently operational site.

Table 2. Experimental Design

#	Orthogonal array			Usability	Branding	Speed
1	1	1	1	Current	Known	Fast
2	1	2	2	Current	Not known	Slow
3	2	1	2	Improved	Known	Slow
4	2	2	1	Improved	Not known	Fast

User Evaluation

In total sixteen users were recruited for the study. All were relatively experienced in Internet use. The users were asked to interact with two of the sites and were then asked to evaluate each against a set of criteria such as

- The extent to which they would return to the site
- The degree to which they would trust the company
- Whether they would recommend it to others
- Whether they would expect a high level of service form the company

In total a score for each user's assessment of each site was obtained. These are presented in Table 3 where the maximum possible score amounts to 54.

Table 3. Data Within the Orthogonal Array

	IS Factors $L_4(2^3)$			Data							
Expt.	A	B	C								
1	1	1	1	33	25	33	39	42	38	48	36
2	1	2	2	16	42	29	32	20	41	28	30
3	2	1	2	17	41	48	42	49	47	47	23
4	2	2	1	48	37	38	35	41	42	42	44

Analysis

In Table 4 we present the specification of the 'favoured' website – one with improved usability, known branding and fast speed – together with the contribution that these three factors made to user assessment.

Table 4. Optimum Website

Factor	Opt. Level description	Opt level	Contribution
Usability	Improved	2	3.406
Branding	Known	1	1.343.
Speed	Fast	1	2.156
Total contribution from all factors			6.905
	Current grand average of performance		30.656
	Expected results at optimum condition		43.561

By undertaking ANOVA it is possible to determine the percentage contribution that each of these factors made to overall user assessment. From Table 5 it is clear that perceived usability was by far the most significant factor. Only 14% of the variability in the whole system is accounted by the three factors but this is not a concern. Indeed we would expect the 'other/error' term to be high as it takes account of all other aspects of variability within the system such as variability in the critical judgement of users, and other differences between users (age, gender etc) which could be examined separately in a more extensive study.

Table 5. ANOVA

Factor	DOF	Sum of squares (S)	Variance (V)	F _Ratio	Pure Sum (S')	Percent (%)
Usability	1	371.281	371.281	4.999	297.017	**11.177**
Branding	1	57.781	57.781	0.778	0	0
Speed	1	148.781	148.781	2.003	74.517	2.804
Other / error	28	2079.374	74.263	---------	---------	86.019
Total	31	2657.218	---------	---------	---------	100

Fcrit 1,12(0.05) = 4.75 Fcrit 1,12(0.10) = 9.33

The study itself indicated (Table 5) that the perceived usability was by far the most important factor in the user assessment of the website – far more important than both branding or download time. In the new media arena where branding issues are often seen as paramount the results are rather challenging. Figure 3 and Figure 4 below show the variance in the whole system and the relative importance of the controlled factors for this case study.

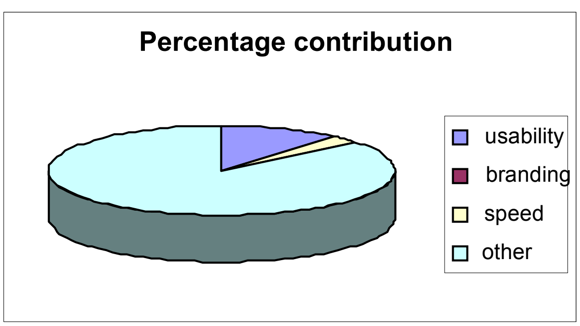

Figure 3. Variance in the Whole System

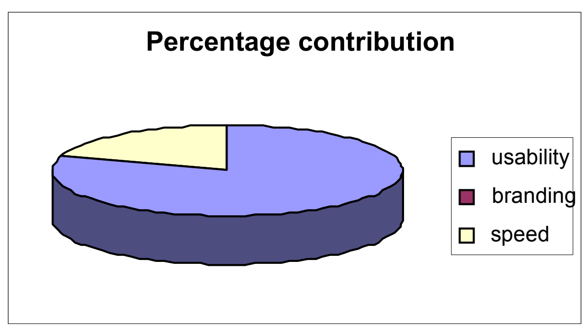

Figure 4. Relative Importance of the Controlled Factors

Conclusion

The research reported in this paper concludes that both the Taguchi and C-Assure Matrices can be very useful at different levels of analysis in ensuring that the Empirical and Holistic Heuristic Evaluation (H^2) is conducted efficiently. We conclude that the two methods can be deployed as part of a holistic usability evaluation toolkit (C-Assure-Taguchi Platform) that is cost effective as well as including the consideration of the more complex, "messy", socio-technical and socio-psychology related issues surrounding the adoption of Information Systems, their various impacts on the organisation and the stakeholders.

Acknowledgement

The authors would like to acknowledge the financial support provided by the UK Engineering and Physical Sciences Research Council (EPSRC) Grant Ref: [GR/R08025).

References

Anthony R A, Planning and control systems: a framework for analysis, Harvard University, Boston 1965.

Badii A, On-line point-of-click web usability mining with PopEval_MB, WebEval_AB and the C-Assure methodology, Proc. AMCIS2000, UCLB, August 2000.

Badii A, Man-machine mutuality and semiotics: evolving web pages as re-adaptive social stereotypes with attitude, Proc. 2nd EnCKompass, Workshop Trinity College Dublin, September 2000.

Dix A, Finlay J, Abowd G, Beale R, Human-Computer Interaction, Prentice Hall, 1993.

Draper S W, Brown M I, Henderson F P, McAteer E, Integrative evaluation: an emerging role for classroom studies of CAL, Computers & Education 26, 1-3, 1996.

Butler Cox Foundation, Information Technology: value for money, CSC Index, Bloomsbury Square, London 1986.

Dunckley L, Smith A and Howard D, Designing for shared interfaces with diverse user groups, in Sasse M A, and Johnson C (Eds.), Human Computer Interaction – INTERACT 99, pp. 630-636, Chapman and Hall, 1999.

Eason K, IT *and Organisational Change*, Taylor and Francis, 1988.

Gibbs W W, Software's chronic crisis, Scientific American, 27 (3), pp. 72-81, 1994.

Hirschheim R and Smithson S, Evaluation of Information Systems: a critical assessment, in Willcocks, L and Lester S. (eds.), Beyond the Productivity Paradox, John Wiley, Chichester 1999.

Lyytinen K, Expectation failure concept and systems analysts' view of information systems failures: results of an exploratory study, North Holland Information and Management, Vol. 14, pp. 45-56. 1998.

Nielson J, Usability Inspection Methods, John Wiley 1994.

Roy, R., QUALITEK-4: Software for automatic design of experiment using Taguchi approach, available from Ranjit K Roy, NUTEK Inc. 30600 Telegraph Road. Suite 2230, Birmingham, Michigan 48025, demonstration version available on the Internet at www.rkroy.com 1998.

Shneiderman B, The Future of interactive systems and the emergence of direct manipulation, Journal of Behaviour & IT 1,3. 1982.

Smith A, Human Computer Factors: a study of users and information systems*,* Mc-Graw Hill, 1997.

Smith A. and Dunckley, L, Using the LUCID method to optimise the acceptability of shared interfaces, Interacting *with Computers*. **9** (3). pp. 335-344, 1998.

Smith A., Dunckley, L., Burkhardt, D., Murkett, A. Eason, K., and Church, J., Developing the optimum help system using the LUCID method, in S. Howard. J. Hammond. G. Lindgaard (eds.) Human-Computer Interaction, Proceedings of INTERACT 97, pp. 307-314, Chapman and Hall 1997.

Smith A. and Dunckley, L., Towards a total quality interface - the application of TQM techniques to interface design, in Sasse M A, Cunningham R J and Winder R L (Eds.), People & Computers X1, Proceedings of HCI'96 (Eds.), pp. 3-17, Springer, 1996.

Taguchi G, Introduction to Quality Engineering, Asian Productivity Organisation, 1986.

Whyte G and Bythway A, Factors affecting information systems' success, International Journal of Service Industry Management, 7 (1), pp.74-93, 1996.